



1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/786,389C

DATE: 10/21/2003

TIME: 17:58:06

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\10212003\I786389C.raw

RECEIVED
OCT 27 2003
TECH CENTER 1600/2900

4 <110> APPLICANT: Japan Science and Technology Corporation
6 <120> TITLE OF INVENTION: Amino Acid Transporter And Gene Thereof
8 <130> FILE REFERENCE: PC901338
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/786,389C
C--> 11 <141> CURRENT FILING DATE: 2001-07-18
13 <160> NUMBER OF SEQ ID NOS: 27
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 4539
18 <212> TYPE: DNA
19 <213> ORGANISM: Homo sapiens
21 <220> FEATURE:
22 <221> NAME/KEY: 5'UTR
23 <222> LOCATION: (1)..(65)
25 <220> FEATURE:
26 <221> NAME/KEY: CDS
27 <222> LOCATION: (66)..(1589)
29 <220> FEATURE:
30 <221> NAME/KEY: 3'UTR
31 <222> LOCATION: (1590)..(4474)
33 <400> SEQUENCE: 1
34 cggcgcgcac actgctcgtt gggccgcggc tcccggtgtt cccaggcccg gccggtgcgc 60
36 agagc atg gcg ggt gcg ggc ccg aag cgg cgc gcg cta gcg gcg ccg gcg 110
37 Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
38 1 5 10 15
40 gcc gag gag aag gaa gag gcg cgg gag aag atg ctg gcc gcc aag agc 158
41 Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys Ser
42 20 25 30
44 gcg gac ggc tcg gcg ccg gca ggc gag ggc gag ggc gtg acc ctg cag 206
45 Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr Leu Gln
46 35 40 45
48 cgg aac atc acg ctg ctc aac ggc gtg gcc atc atc gtg ggg acc att 254
49 Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val Gly Thr Ile
50 50 55 60
52 atc ggc tcg ggc atc ttc gtg acg ccc acg ggc gtg ctc aag gag gca 302
53 Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val Leu Lys Glu Ala
54 65 70 75
56 ggc tcg ccg ggg ctg gcg ctg gtg tgg gcc gcg tgc ggc gtc ttc 350
57 Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala Ala Cys Gly Val Phe
58 80 85 90 95
60 tcc atc gtg ggc gcg ctc tgc tac gcg gag ctc gcc acc acc atc tcc 398
61 Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu Leu Gly Thr Thr Ile Ser
62 100 105 110
64 aaa tcg ggc ggc gac tac gcc tac atg ctg gag gtc tac ggc tcg ctg 446

ENTERED

RAW SEQUENCE LISTING

DATE: 10/21/2003

PATENT APPLICATION: US/09/786,389C

TIME: 17:58:06

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\10212003\I786389C.raw

```

65 Lys Ser Gly Gly Asp Tyr Ala Tyr Met Leu Glu Val Tyr Gly Ser Leu
66      115      120      125
68 ccc gcc ttc ctc aag ctc tgg atc gag ctg ctc atc atc cgg cct tca 494
69 Pro Ala Phe Leu Lys Leu Trp Ile Glu Leu Leu Ile Ile Arg Pro Ser
70      130      135      140
72 tcg cag tac atc gtg gcc ctg gtc ttc gcc acc tac ctg ctc aag ccg 542
73 Ser Gln Tyr Ile Val Ala Leu Val Phe Ala Thr Tyr Leu Leu Lys Pro
74      145      150      155
76 ctc ttc ccc acc tgc ccg gtg ccc gag gag gca gcc aag ctc gtg gcc 590
77 Leu Phe Pro Thr Cys Pro Val Pro Glu Glu Ala Ala Lys Leu Val Ala
78 160      165      170      175
80 tgc ctc tgc gtg ctg ctg ctc acg gcc gtg aac tgc tac agc gtg aag 638
81 Cys Leu Cys Val Leu Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys
82      180      185      190
84 gcc gcc acc cgg gtc cag gat gcc ttt gcc gcc gcc aag ctc ctg gcc 686
85 Ala Ala Thr Arg Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala
86      195      200      205
88 ctg gcc ctg atc atc ctg ctg ggc ttc gtc cag atc ggg aag ggt gat 734
89 Leu Ala Leu Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp
90      210      215      220
92 gtg tcc aat cta gat ccc aac ttc tca ttt gaa ggc acc aaa ctg gat 782
93 Val Ser Asn Leu Asp Pro Asn Phe Ser Phe Glu Gly Thr Lys Leu Asp
94      225      230      235
96 gtg ggg aac att gtg ctg gca tta tac agc ggc ctc ttt gcc tat gga 830
97 Val Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
98 240      245      250      255
100 gga tgg aat tac ttg aat ttc gtc aca gag gaa atg atc aac ccc tac 878
101 Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro Tyr
102      260      265      270
104 aga aac ctg ccc ctg gcc atc atc atc tcc ctg ccc atc gtg acg ctg 926
105 Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val Thr Leu
106      275      280      285
108 gtg tac gtg ctg acc aac ctg gcc tac ttc acc acc ctg tcc acc gag 974
109 Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu Ser Thr Glu
110      290      295      300
112 cag atg ctg tcg tcc gag gcc gtg gcc gtg gac ttc ggg aac tat cac 1022
113 Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe Gly Asn Tyr His
114      305      310      315
116 ctg ggc gtc atg tcc tgg atc atc ccc gtc ttc gtg ggc ctg tcc tgc 1070
117 Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe Val Gly Leu Ser Cys
118 320      325      330      335
120 ttc ggc tcc gtc aat ggg tcc ctg ttc aca tcc tcc agg ctc ttc ttc 1118
121 Phe Gly Ser Val Asn Gly Ser Leu Phe Thr Ser Ser Arg Leu Phe Phe
122      340      345      350
124 gtg ggg tcc cgg gaa ggc cac ctg ccc tcc atc ctc tcc atg atc cac 1166
125 Val Gly Ser Arg Glu Gly His Leu Pro Ser Ile Leu Ser Met Ile His
126      355      360      365
128 cca cag ctc ctc acc ccc gtg ccg tcc ctc gtg ttc acg tgt gtg atg 1214
129 Pro Gln Leu Leu Thr Pro Val Pro Ser Leu Val Phe Thr Cys Val Met

```

RAW SEQUENCE LISTING

DATE: 10/21/2003

PATENT APPLICATION: US/09/786,389C

TIME: 17:58:06

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\10212003\I786389C.raw

```

130          370          375          380
132 acg ctg ctc tac gcc ttc tcc aag gac atc ttc tcc gtc atc aac ttc 1262
133 Thr Leu Leu Tyr Ala Phe Ser Lys Asp Ile Phe Ser Val Ile Asn Phe
134          385          390          395
136 ttc agc ttc ttc aac tgg ctc tgc gtg gcc ctg gcc atc atc ggc atg 1310
137 Phe Ser Phe Phe Asn Trp Leu Cys Val Ala Leu Ala Ile Ile Gly Met
138 400          405          410          415
140 atc tgg ctg cgc cac aga aag cct gag ctt gag cgg ccc atc aag gtg 1358
141 Ile Trp Leu Arg His Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val
142          420          425          430
144 aac ctg gcc ctg cct gtg ttc ttc atc ctg gcc tgc ctc ttc ctg atc 1406
145 Asn Leu Ala Leu Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile
146          435          440          445
148 gcc gtc tcc ttc tgg aag aca ccc gtg gag tgt ggc atc ggc ttc acc 1454
149 Ala Val Ser Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr
150          450          455          460
152 atc atc ctc agc ggg ctg ccc gtc tac ttc ttc ggg gtc tgg tgg aaa 1502
153 Ile Ile Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys
154          465          470          475
156 aac aag ccc aag tgg ctc ctc cag ggc atc ttc tcc acg acc gtc ctg 1550
157 Asn Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
158 480          485          490          495
160 tgt cag aag ctc atg cag gtg gtc ccc cag gag aca tag ccaggaggcc 1599
161 Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
162          500          505
164 gagtggtgc cggaggagca tgcgcagagg ccagttaaag tagatcacct cctcgaaccc 1659
166 actcgggttc cccgcaaccc acagctcagc tgcccatccc agtccctcgc cgtccctccc 1719
168 aggtcgggca gtggaggctg ctgtgaaac tctggtacga atctcatccc tcaactgagg 1779
170 gccagggacc caggtgtgcc tgtgtcctg cccaggagca gcttttggtc tccttggggc 1839
172 ctttttcctt tccctccttt gtttacttat atatatattt tttttaaaact taaatttttg 1899
174 gtcaacttga caccactaag atgatttttt aaggagctgg gggaaggcag gagccttcct 1959
176 ttctcctgcc ccaagggcc agaccctggg caaacagagc tactgagact tggaacctca 2019
178 ttgctacgac agacttgac tgaagccgga cagctgccca gacacatggg cttgtgacat 2079
180 tcgtgaaaac caaccctgtg ggcttatgtc tctgccttag ggtttgaga gtggaaactc 2139
182 agccgtaggg tggcactggg aggggggtgg ggatctgggc aaggtgggtg attcctccca 2199
184 ggaggtgctt gagggcccga tggactcctg accataatcc tagccccgag acaccatcct 2259
186 gagccaggga acagcccag ggttgggggg tgccggcatc tcccctagct caccaggcct 2319
188 ggcctctggg cagtgtggcc tcttggtat ttctgttcca gttttggagg ctgagttctg 2379
190 gttcatgcag acaaagccct gtccttcagt cttctagaaa cagagacaag aaaggcagac 2439
192 acaccgcggc caggcaccca tgtgggcgcc caccctgggc tccacacagc agtgtcccct 2499
194 gcccagagg tgcagctac cctcagcctc caatgcattg gcctctgtac cgcccggcag 2559
196 ccccttctgg ccggtgctgg gttccactc ccggcctagg cacctccccg ctctccctgt 2619
198 cacgctcatg tcctgtcctg gtctgatgc ccgttgtcta ggagacagag ccaagcactg 2679
200 ctacgtctc tgccgcctgc gtttgagggc ccctgggctc taccacagtc cccaccgcc 2739
202 tgcagagagg gaactagggc accccttgtt tctgttgttc ccgtgaattt ttttcgctat 2799
204 gggaggcagc cgaggcctgg ccaatgcggc ccactttcct gagctgtcgc tgcctccatg 2859
206 gcagcagcca aggaccccca gaacaagaag acccccccgc aggatccctc ctgagctcgg 2919
208 ggggctctgc cttctcaggc cccgggcttc cttctcccc agccagaggg ggagccaagt 2979
210 ggtccagcgt cactccagtg ctcagctgtg gctggaggag ctggcctgtg gcacagccct 3039

```

RAW SEQUENCE LISTING

DATE: 10/21/2003

PATENT APPLICATION: US/09/786,389C

TIME: 17:58:06

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\10212003\I786389C.raw

```

212 gagtgtccca agccgggagc caacgaagcc ggacacggct tcaactgacca gcggtctgtc 3099
214 aagccgcaag ctctcagcaa gtgcccagtg gagcctgccg cccccacctg ggcaccggga 3159
216 cccctcacc atccagtggtg cccggagaaa cctgatgaac agtttgggga ctgaggacca 3219
218 gatgtccgtc tctcttgctt gaggaatgaa gacctttatt caccctgtcc ccgttgcttc 3279
220 ccgctgcaca tggacagact tcacagcgctc tgctcatagg acctgcatcc ttcctgggga 3339
222 cgaattccac tcgtccaagg gacagccac ggtctggagg ccgaggacca ccagcaggca 3399
224 ggtggactga ctgtgttggg caagacctct tccctctggg cctgttctct tggctgcaaa 3459
226 taaggacagc agctgggtgc ccacctgcct ggtgcattgc tgtgtgaatc caggaggcag 3519
228 tggacatcgt aggcagccac ggccccgggt ccaggagaag tgctccctgg aggcacgcac 3579
230 cactgcttcc cactggggcc ggcgggggcc acgcacgacg tcagcctctt accttcccgc 3639
232 ctgggttagg ggtcctcggg atgccgttct gttccaacct cctgctctgg gaggtggaca 3699
234 tgcctcaagg atacaggag ccggcgccct ctgcacggca cgcacttgcc tgttggtgc 3759
236 tgcggctgtg ggcgagcatg ggggctgcc gcgctctgtg tggaaagtag ctgctagtga 3819
238 aatggctggg gccgctgggg tccgtcttca cactgcgcag gtctcttctg ggcgtctgag 3879
240 ctgggggtgg agctcctccg cagaaggttg gtggggggtc cagtctgtga tccttggtgc 3939
242 tgtgtgcccc actccagcct ggggacccca cttcagaagg taggggccgt gtcccgcggt 3999
244 gctgactgag gcctgcttcc ccctccccct cctgctgtgc tgggaattcca cagggaccag 4059
246 ggccaccgca ggggactgtc tcagaagact tgatttttcc gtcccttttt ctccacactc 4119
248 cactgacaaa cgtccccagc ggtttccact tgtgggcttc aggtgttttc aagcacaacc 4179
250 caccacaaca agcaagtgca ttttcagtcg ttgtgctttt ttgttttgtg ctaacgtctt 4239
252 actaatttaa agatgctgtc ggcaccatgt ttattttatt ccagtgggtc tgctcagcct 4299
254 tgctgctctg cgtggcgagc gtgccatgcc tgctccctgt ctgtgtccca gccacgcagg 4359
256 gccatccact gtgacgtcgg ccgaccaggc tggacaccct ctgccgagta atgacgtgtg 4419
258 tggctgggac cttctttatt ctgtgttaat ggctaacctg ttacactggg ctgggttggg 4479
260 taggggtgtc tggctttttt gtgggggttt tatttttaaa gaaacactca atcatcctag 4539

```

263 <210> SEQ ID NO: 2

264 <211> LENGTH: 507

265 <212> TYPE: PRT

266 <213> ORGANISM: Homo sapiens

268 <400> SEQUENCE: 2

```

269 Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala Ala
270   1               5               10              15
272 Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys Ser Ala
273           20           25           30
275 Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr Leu Gln Arg
276       35           40           45
278 Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val Gly Thr Ile Ile
279   50           55           60
281 Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val Leu Lys Glu Ala Gly
282  65           70           75           80
284 Ser Pro Gly Leu Ala Leu Val Val Trp Ala Ala Cys Gly Val Phe Ser
285           85           90           95
287 Ile Val Gly Ala Leu Cys Tyr Ala Glu Leu Gly Thr Thr Ile Ser Lys
288       100          105          110
290 Ser Gly Gly Asp Tyr Ala Tyr Met Leu Glu Val Tyr Gly Ser Leu Pro
291   115          120          125
293 Ala Phe Leu Lys Leu Trp Ile Glu Leu Leu Ile Ile Arg Pro Ser Ser
294   130          135          140
296 Gln Tyr Ile Val Ala Leu Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu

```

RAW SEQUENCE LISTING

DATE: 10/21/2003

PATENT APPLICATION: US/09/786,389C

TIME: 17:58:06

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\10212003\I786389C.raw

```

297 145          150          155          160
299 Phe Pro Thr Cys Pro Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys
300          165          170          175
302 Leu Cys Val Leu Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala
303          180          185          190
305 Ala Thr Arg Val Gln Asp Ala Phe Ala Ala Lys Leu Leu Ala Leu
306          195          200          205
308 Ala Leu Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val
309          210          215          220
311 Ser Asn Leu Asp Pro Asn Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
312 225          230          235          240
314 Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly Gly
315          245          250          255
317 Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro Tyr Arg
318          260          265          270
320 Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val Thr Leu Val
321          275          280          285
323 Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu Ser Thr Glu Gln
324          290          295          300
326 Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe Gly Asn Tyr His Leu
327 305          310          315          320
329 Gly Val Met Ser Trp Ile Ile Pro Val Phe Val Gly Leu Ser Cys Phe
330          325          330          335
332 Gly Ser Val Asn Gly Ser Leu Phe Thr Ser Ser Arg Leu Phe Phe Val
333          340          345          350
335 Gly Ser Arg Glu Gly His Leu Pro Ser Ile Leu Ser Met Ile His Pro
336          355          360          365
338 Gln Leu Leu Thr Pro Val Pro Ser Leu Val Phe Thr Cys Val Met Thr
339          370          375          380
341 Leu Leu Tyr Ala Phe Ser Lys Asp Ile Phe Ser Val Ile Asn Phe Phe
342 385          390          395          400
344 Ser Phe Phe Asn Trp Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile
345          405          410          415
347 Trp Leu Arg His Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn
348          420          425          430
350 Leu Ala Leu Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala
351          435          440          445
353 Val Ser Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile
354          450          455          460
356 Ile Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
357 465          470          475          480
359 Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu Cys
360          485          490          495
362 Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
363          500          505
366 <210> SEQ ID NO: 3
367 <211> LENGTH: 3455
368 <212> TYPE: DNA
369 <213> ORGANISM: Rat

```

VERIFICATION SUMMARY

DATE: 10/21/2003

PATENT APPLICATION: US/09/786,389C

TIME: 17:58:07

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\10212003\I786389C.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:837 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:1569 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ ID#:26
L:1569 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:624
M:341 Repeated in SeqNo=26
L:1660 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ ID#:27
L:1660 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:48
M:341 Repeated in SeqNo=27